GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 19, 2005, 15:35:11 ; Search time 43.7895 Seconds
(without alignments)
1413.163 Million cell updates/sec

Title: US-09-904-954-4_COPY_1_160
Perfect score: 879
Sequence: 1 MRLLPLLRTVLWAAFLGSPL.....SGQCLRLQVSVCCKERKSES 160

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	879	100.0	201	2	AAR71482	Aar71482 Human hek

ALIGNMENTS

RESULT 1
AAR71482
ID AAR71482 standard; protein; 201 AA.
XX
AC AAR71482;
XX
DT 25-MAR-2003 (revised)
DT 03-OCT-1995 (first entry)
XX
DE Human hek-L protein.
XX
KW Ligand; cell surface; tyrosine kinase receptor; tumorigenesis; immunogen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .22
FT /note= "signal peptide"
FT Protein 23. .201
XX
PN WO9506065-A1.
XX
PD 02-MAR-1995.
XX
PF 17-AUG-1994; 94WO-US009282.
XX
PR 20-AUG-1993; 93US-00109745.
PR 30-AUG-1993; 93US-00114426.
PR 03-DEC-1993; 93US-00161132.
PR 09-MAY-1994; 94US-00240124.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Beckmann MP, Cerretti DP;
XX
DR WPI; 1995-106811/14.
DR N-PSDB; AAQ85888.
XX
PT New isolated DNA encoding hek-L protein or its fusion products - useful
PT as assay reagent or for carrying therapeutic and diagnostic compounds to
PT leukaemia cells.
XX
PS Claim 21; Page 38; 45pp; English.
XX
CC The sequence is that of a novel protein designated hek-L, a protein that
CC can bind hek (a cell surface receptor tyrosine kinase). Hek-L is the
CC first known ligand for hek and can be used to study cellular processes
CC regulated by hek (which may be involved in tumorigenesis). It is also an
CC immunogen for antibody production, as a reagent for detecting hek or hek-
CC L in in vitro assays, to determine binding of hek proteins, to purify hek
CC proteins, and to carry diagnostic or cytotoxic agents to particular
CC leukaemia cells that express the hek antigen. Hek-L also binds the elk
CC tyrosine kinase receptors. See also AAR71481. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 201 AA;

Query Match 100.0%; Score 879; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.4e-93;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSES 160
              ||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSES 160
```

Search completed: April 19, 2005, 23:13:33
Job time : 43.7895 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 09:30:04 ; Search time 3056 Seconds
(without alignments)
8183.329 Million cell updates/sec

Title: US-09-904-954-1_COPY_140_796
Perfect score: 657
Sequence: 1 CTGCTGGCCCAAGGGCCCGG.....TCATGACGTTCTTGGCCTCC 657

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	657	100.0	1741	3	CR607482	CR607482 full-leng
2	645	98.2	871	5	BX419695	BX419695 BX419695
3	622.6	94.8	1100	1	AL533153	AL533153 AL533153
4	598.4	91.1	726	1	AL527972	AL527972 AL527972
5	587.2	89.4	1629	3	CR597504	CR597504 full-leng

ALIGNMENTS

RESULT 1
CR607482
LOCUS CR607482 1741 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DN003YL19 of Adult brain of Homo sapiens (human).
ACCESSION CR607482
VERSION CR607482.1 GI:50488289
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1741)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 1741)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .1741
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN003YL19"
/tissue_type="Adult brain"
/plasmid="pCMVSPORT_6"
ORIGIN

Query Match 100.0%; Score 657; DB 3; Length 1741;
Best Local Similarity 100.0%; Pred. No. 6.2e-143;
Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTGCTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGC 181

Qy         61 TCCAACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTG 241

Qy        121 GATATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GATATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCC 301

Qy        181 GGAGGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GGAGGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGC 361

Qy        241 AACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCC 421

Qy        301 ATCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCAC 481

Qy        361 GCCGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GCCGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTG 541

Qy        421 AGGATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AGGATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCC 601

Qy        481 ACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAG 661

Qy        541 GGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGG 721

Qy        601 GAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 657
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 778
```

```
RESULT 2
BX419695
LOCUS       BX419695                871 bp    mRNA    linear   EST 01-MAY-2004
DEFINITION  BX419695 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF019YB24 5-PRIME, mRNA sequence.
ACCESSION   BX419695
VERSION     BX419695.2  GI:46934231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30654816.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 497.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF019DA12QP1&c=497.f.
FEATURES             Location/Qualifiers
     source          1. .871
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF019YB24"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             98.2%;  Score 645;  DB 5;  Length 871;
  Best Local Similarity   99.8%;  Pred. No. 3.7e-140;
  Matches  656;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CTGCTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGC 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        114 CTGCTGGCCCAAGGGCCCGGA-GGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGC 172

Qy         61 TCCAACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TCCAACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTG 232

Qy        121 GATATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GATATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCC 292

Qy        181 GGAGGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GGAGGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGC 352

Qy        241 AACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 AACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCC 412

Qy        301 ATCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCAC 472

Qy        361 GCCGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 GCCGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTG 532

Qy        421 AGGATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AGGATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCC 592

Qy        481 ACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 ACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAG 652

Qy        541 GGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGG 712

Qy        601 GAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 657
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 GAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 769

RESULT 3
AL533153
LOCUS       AL533153                1100 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL533153 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN003YL19 5-PRIME, mRNA sequence.
ACCESSION   AL533153
VERSION     AL533153.3  GI:45708058
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1100)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31070985.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 497.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN003CF10QP1&c=497.f.
FEATURES             Location/Qualifiers
     source          1. .1100
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN003YL19"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             94.8%;  Score 622.6;  DB 1;  Length 1100;
  Best Local Similarity   99.4%;  Pred. No. 6.5e-135;
  Matches  655;  Conservative    1;  Mismatches    0;  Indels    3;  Gaps    3;

Qy          1 CTGCTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGC 60
```

```
Db        122 CTGCTGGCCCAAGGGCCCGGA-GGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGC 180

Qy         61 TCCAACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTG 120

Db        181 TCCAACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTG 240

Qy        121 GATATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCC 180

Db        241 GATATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCC 300

Qy        181 GGAGGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGC 240

Db        301 GGAGGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGC 360

Qy        241 AACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCC 300

Db        361 AACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCC 420

Qy        301 ATCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCT-GGGCTACGAGTTCCA 359

Db        421 ATCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCSCCTTCTCTCTGGGGCTACGAGTTCCA 480

Qy        360 CGCCGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCT 419

Db        481 CGCCGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCT 540

Qy        420 GAGGATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCC-GGGGAGAAGCCGGTCC 478

Db        541 GAGGATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGGAGAAGCCGGTCC 600

Qy        479 CCACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTG 538

Db        601 CCACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTG 660

Qy        539 AGGGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAAC 598

Db        661 AGGGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAAC 720

Qy        599 GGGAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 657

Db        721 GGGAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 779
```

```
RESULT 4
AL527972
LOCUS       AL527972                726 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL527972 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC027YP19 5-PRIME, mRNA sequence.
ACCESSION   AL527972
VERSION     AL527972.3  GI:45703058
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31065823.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            497.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DC027CH10QP1&c=497.f.
```

```
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC027YP19"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

```
  Query Match             91.1%;  Score 598.4;  DB 1;  Length 726;
  Best Local Similarity   91.0%;  Pred. No. 2.8e-129;
  Matches  595;  Conservative   41;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          1 CTGCTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGC 60

Db         69 CTGCTGGCCCAAGGGCCCGGA-GGGCGAAGGGAAACMGGCATGCGGTGTACTGGAACARC 127

Qy         61 TCCAACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTG 120

Db        128 WCCAACCAGCACCTGCGGCGAGAGGGCWACAMMAAGMAGGTGAACGTGAACGACTATMTG 187

Qy        121 GATATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCC 180

Db        188 GATAWWWAAAGCACGMAMTAAAAMAGCWCGGGGGTRGGCCCCGGGGCGGGAMCGGGGCCC 247

Qy        181 GGAGGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGC 240

Db        248 GGAGGCGGGGCARARCAGTACGTGCTGTACATGGWGAGCMGCAACGGCWACMGCACCTGC 307

Qy        241 AACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCC 300

Db        308 AACGCCAGCAAGGGCTTMAAGCGCTGGGAGTGCAACCGGCCGCACGCMCCGCAMAGCCCC 367

Qy        301 ATCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCAC 360

Db        368 ATMAAGTTCTCGGAGAAGTTCCAGCGCWAAARCGCCTTCWMTCTGGGCWACGARTWCCAC 427

Qy        361 GCCGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTG 420

Db        428 GCMGGCCACGAGTACTACWAMATCTCCACGCCCACTCAMAACCTGCACTGGAAGTGTCTG 487

Qy        421 AGGATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCC 480

Db        488 ARGATGAAGGTGTTCGTCTGCAGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCC 547

Qy        481 ACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAG 540

Db        548 ACTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATMAACGTGCTGGAAGACTTTGAG 607

Qy        541 GGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGG 600

Db        608 GGAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGG 667

Qy        601 GAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCC 654

Db        668 GAACACCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCMCGACGTATCTTGGC 721
```

```
RESULT 5
CR597504
LOCUS       CR597504                1629 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI026YI24 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR597504
VERSION     CR597504.1  GI:50478311
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 1629)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1629)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
FEATURES division of Invitrogen.
source Location/Qualifiers
1. .1629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI026YI24"
/tissue_type="Placenta Cot 25-normalized"
ORIGIN /plasmid="pCMVSPORT_6"

Search completed: April 19, 2005, 21:57:08
Job time : 3057.14 secs

Query Match 89.4%; Score 587.2; DB 3; Length 1629;
Best Local Similarity 99.5%; Pred. No. 1.3e-126;
Matches 589; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        66 CCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGATAT 125
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        77 CCACAGCCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGATAT 125

Qy       126 TTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGAGG 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       137 TTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGAGG 185

Qy       186 CGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAACGC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       197 CGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAACGC 245

Qy       246 CAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATCAA 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       257 CAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATCAA 305

Qy       306 GTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCACGCCGG 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       317 GTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCACGCCGG 365

Qy       366 CCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAGGAT 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       377 CCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAGGAT 425

Qy       426 GAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACTCT 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       437 GAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACTCT 485

Qy       486 CCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGAGA 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       497 CCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGAGA 545

Qy       546 GAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAACA 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       557 GAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAACA 605

Qy       606 CCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 657
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       617 CCTGCCCCTGGCCGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGCCTCC 668
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 09:30:04 ; Search time 3083.91 Seconds
(without alignments)
8183.329 Million cell updates/sec

Title: US-09-904-954-1_COPY_83_745
Perfect score: 663
Sequence: 1 ATGGCGGCGGCTCCGCTGCT.....CCAGCCCCAAACGGGAACAC 663

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	663	100.0	1741	3	CR607482	CR607482 full-leng
2	651	98.2	871	5	BX419695	BX419695 BX419695
3	628.6	94.8	1100	1	AL533153	AL533153 AL533153
4	603.8	91.1	726	1	AL527972	AL527972 AL527972

ALIGNMENTS

RESULT 1
CR607482
LOCUS CR607482 1741 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DN003YL19 of Adult brain of Homo sapiens (human).
ACCESSION CR607482
VERSION CR607482.1 GI:50488289
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1741)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 1741)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .1741
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN003YL19"
/tissue_type="Adult brain"
/plasmid="pCMVSPORT_6"
ORIGIN

Query Match 100.0%; Score 663; DB 3; Length 1741;
Best Local Similarity 100.0%; Pred. No. 6.4e-137;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGCGGCTCCGCTGCTGCTGCTGCTGCTGCTCGTGCCCGTGCCGCTGCTGCCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ATGGCGGCGGCTCCGCTGCTGCTGCTGCTGCTCGTGCCCGTGCCGCTGCTGCCGCTG 124

Qy         61 CTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGCTCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGCTCC 184

Qy        121 AACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 AACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGAT 244

Qy        181 ATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGA 304

Qy        241 GGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAAC 364

Qy        301 GCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 GCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATC 424

Qy        361 AAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCACGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 AAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCACGCC 484

Qy        421 GGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 GGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAGG 544

Qy        481 ATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 ATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACT 604

Qy        541 CTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 CTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGA 664

Qy        601 GAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 GAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAA 724
```

```
Qy          661 CAC 663
                ||| 
Db          725 CAC 727

RESULT 2
BX419695
LOCUS       BX419695                871 bp    mRNA    linear   EST 01-MAY-2004
DEFINITION  BX419695 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF019YB24 5-PRIME, mRNA sequence.
ACCESSION   BX419695
VERSION     BX419695.2  GI:46934231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30654816.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 497.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF019DA12QP1&c=497.f.
FEATURES
     source          Location/Qualifiers
                     1. .871
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF019YB24"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match                98.2%;  Score 651;  DB 5;  Length 871;
  Best Local Similarity      99.8%;  Pred. No. 2.8e-134;
  Matches  662;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 ATGGCGGCGGCTCCGCTGCTGCTGCTGCTGCTCGTGCCCGTGCCGCTGCTGCCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 ATGGCGGCGGCTCCGCTGCTGCTGCTGCTGCTCGTGCCCGTGCCGCTGCTGCCGCTG 116

Qy         61 CTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGCTCC 120
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        117 CTGGCCCAAGGGCCCGGA-GGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGCTCC 175

Qy        121 AACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 AACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGAT 235

Qy        181 ATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 ATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGA 295

Qy        241 GGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 GGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAAC 355

Qy        301 GCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 GCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATC 415

Qy        361 AAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTCTGGGCTACGAGTTCCACGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 AAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTCTGGGCTACGAGTTCCACGCC 475

Qy        421 GGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 GGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAGG 535

Qy        481 ATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 ATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACT 595

Qy        541 CTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        596 CTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGA 655

Qy        601 GAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        656 GAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAA 715

Qy        661 CAC 663
                ||| 
Db        716 CAC 718

RESULT 3
AL533153
LOCUS       AL533153                1100 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL533153 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN003YL19 5-PRIME, mRNA sequence.
ACCESSION   AL533153
VERSION     AL533153.3  GI:45708058
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1100)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31070985.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 497.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN003CF10QP1&c=497.f.
FEATURES
     source          Location/Qualifiers
                     1. .1100
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN003YL19"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized "
```

```
ORIGIN

  Query Match             94.8%;  Score 628.6;  DB 1;  Length 1100;
  Best Local Similarity   99.4%;  Pred. No. 2.7e-129;
  Matches  661;  Conservative    1;  Mismatches    0;  Indels    3;  Gaps    3;

Qy          1 ATGGCGGCGGCTCCGCTGCTGCTGCTGCTGCTGCTCGTGCCCGTGCCGCTGCTGCCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ATGGCGGCGGCTCCGCTGCTGCTGCTGCTGCTCGTGCCCGTGCCGCTGCTGCCGCTG 124

Qy         61 CTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGCTCC 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        125 CTGGCCCAAGGGCCCGGA-GGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGCTCC 183

Qy        121 AACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGAT 243

Qy        181 ATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGA 303

Qy        241 GGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAAC 363

Qy        301 GCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATC 423

Qy        361 AAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCT-GGGCTACGAGTTCCACGC 419
              ||||||||||||||||||||||||||||||:|||||||||| ||||||||||||||||||
Db        424 AAGTTCTCGGAGAAGTTCCAGCGCTACAGCSCCTTCTCTCTGGGGCTACGAGTTCCACGC 483

Qy        420 CGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAG 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 CGGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAG 543

Qy        480 GATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCC-GGGGAGAAGCCGGTCCCCA 538
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        544 GATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGGAGAAGCCGGTCCCCA 603

Qy        539 CTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGG 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 CTCTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGG 663

Qy        599 GAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGG 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 GAGAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGG 723

Qy        659 AACAC 663
              |||||
Db        724 AACAC 728


RESULT 4
AL527972
LOCUS       AL527972                726 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL527972 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC027YP19 5-PRIME, mRNA sequence.
ACCESSION   AL527972
VERSION     AL527972.3  GI:45703058
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31065823.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            497.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DC027CH10QP1&c=497.f.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC027YP19"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             91.1%;  Score 603.8;  DB 1;  Length 726;
  Best Local Similarity   89.0%;  Pred. No. 8.6e-124;
  Matches  590;  Conservative   48;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGCGGCTCCGCTGCTGCTGCTGCTGCTGCTCGTGCCCGTGCCGCTGCTGCCGCTG 60
              ||||||||||:|:||:||:||:||:||:|||||||||||:|||||||||||||||||
Db         11 ATGGCGGCGGCWCMGCWGCWGCWGCWGCWGCTGCTCGTGCMCGTGCCGCTGCTGCCAGCT 70

Qy         61 CTGGCCCAAGGGCCCGGAGGGGCGCTGGGAAACCGGCATGCGGTGTACTGGAACAGCTCC 120
                 || | || || | |||||  |||||||:||||||||||||||||||:|:||
Db         71 GCTGGCCCAAGGGCCCGGAGGGCGAAGGGAAACMGGCATGCGGTGTACTGGAACARCWCC 130

Qy        121 AACCAGCACCTGCGGCGAGAGGGCTACACCGTGCAGGTGAACGTGAACGACTATCTGGAT 180
              |||||||||||||||||||||||:|||::  |:||||||||||||||||||:|||||
Db        131 AACCAGCACCTGCGGCGAGAGGGCWACAMMAAGMAGGTGAACGTGAACGACTATMTGGAT 190

Qy        181 ATTTACTGCCCGCACTACAACAGCTCGGGGGTGGGCCCCGGGGCGGGACCGGGGCCCGGA 240
              |:::|  || ||:|:|| ||:|||:|||||||:|||||||||||||||:||||||||||
Db        191 AWWWAAAGCACGMAMTAAAAMAGCWCGGGGGTRGGCCCCGGGGCGGGAMCGGGGCCCGGA 250

Qy        241 GGCGGGGCAGAGCAGTACGTGCTGTACATGGTGAGCCGCAACGGCTACCGCACCTGCAAC 300
              |||||||||:|:|||||||||||||||||||:||||:|||||||:||:||||||||||
Db        251 GGCGGGGCARARCAGTACGTGCTGTACATGGWGAGCMGCAACGGCWACMGCACCTGCAAC 310

Qy        301 GCCAGCCAGGGCTTCAAGCGCTGGGAGTGCAACCGGCCGCACGCCCCGCACAGCCCCATC 360
              ||||||  |||||||:|||||||||||||||||||||||||:|||||:|||||||||:
Db        311 GCCAGCAAGGGCTTMAAGCGCTGGGAGTGCAACCGGCCGCACGCMCCGCAMAGCCCCATM 370

Qy        361 AAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCCTTCTCTCTGGGCTACGAGTTCCACGCC 420
              |||||||||||||||||||||||:| |:|||||||::|||||||:||||:|:||||||:
Db        371 AAGTTCTCGGAGAAGTTCCAGCGCWAAARCGCCTTCWMTCTGGGCWACGARTWCCACGCM 430

Qy        421 GGCCACGAGTACTACTACATCTCCACGCCCACTCACAACCTGCACTGGAAGTGTCTGAGG 480
              |||||||||||||:|:||||||||||||||||:|||||||||||||||||||||||:|
Db        431 GGCCACGAGTACTACWAMATCTCCACGCCCACTCAMAACCTGCACTGGAAGTGTCTGARG 490

Qy        481 ATGAAGGTGTTCGTCTGCTGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACT 540
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        491 ATGAAGGTGTTCGTCTGCAGCGCCTCCACATCGCACTCCGGGGAGAAGCCGGTCCCCACT 550

Qy        541 CTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGA 600
              |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        551 CTCCCCCAGTTCACCATGGGCCCCAATGTGAAGATMAACGTGCTGGAAGACTTTGAGGGA 610

Qy        601 GAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 GAGAACCCTCAGGTGCCCAAGCTTGAGAAGAGCATCAGCGGGACCAGCCCCAAACGGGAA 670
```

```
Qy        661 CAC 663
              |||
Db        671 CAC 673
```

Search completed: April 19, 2005, 21:57:07
Job time : 3085.05 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 19, 2005, 15:35:11 ; Search time 59.9368 Seconds
(without alignments)
1413.163 Million cell updates/sec

Title: US-09-904-954-2_COPY_1_219
Perfect score: 1207
Sequence: 1 MAAAPLLLLLLLVPVPLLPL.....GENPQVPKLEKSISGTSPKR 219

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1207	100.0	238	2	AAR71481	Aar71481 Human hek
2	1207	100.0	238	5	AAG79439	Aag79439 Breast/co
3	1207	100.0	238	6	ABP97190	Abp97190 Tumour-as
4	1207	100.0	238	7	ADN38812	Adn38812 Cancer/an
5	1207	100.0	238	8	ADQ21666	Adq21666 Human sof
6	1172	97.1	234	2	AAR82605	Aar82605 Eph trans

ALIGNMENTS

RESULT 1
AAR71481
ID AAR71481 standard; protein; 238 AA.
XX
AC AAR71481;
XX
DT 25-MAR-2003 (revised)
DT 03-OCT-1995 (first entry)
XX
DE Human hek-L protein.
XX
KW Ligand; cell surface; tyrosine kinase receptor; tumorigenesis; immunogen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .19
FT /note= "signal peptide"
FT Protein 20. .238
XX
PN WO9506065-A1.
XX
PD 02-MAR-1995.
XX
PF 17-AUG-1994; 94WO-US009282.
XX
PR 20-AUG-1993; 93US-00109745.
PR 30-AUG-1993; 93US-00114426.
PR 03-DEC-1993; 93US-00161132.
PR 09-MAY-1994; 94US-00240124.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Beckmann MP, Cerretti DP;
XX
DR WPI; 1995-106811/14.
DR N-PSDB; AAQ85887.
XX
PT New isolated DNA encoding hek-L protein or its fusion products - useful
PT as assay reagent or for carrying therapeutic and diagnostic compounds to
PT leukaemia cells.
XX
PS Claim 21; Page 36; 45pp; English.
XX
CC The sequence is that of a novel protein designated hek-L, a protein that
CC can bind hek (a cell surface receptor tyrosine kinase). Hek-L is the
CC first known ligand for hek and can be used to study cellular processes
CC regulated by hek (which may be involved in tumorigenesis). It is also an
CC immunogen for antibody production, as a reagent for detecting hek or hek-
CC L in in vitro assays, to determine binding of hek proteins, to purify hek
CC proteins, and to carry diagnostic or cytotoxic agents to particular
CC leukaemia cells that express the hek antigen. Hek-L also binds the elk
CC tyrosine kinase receptors. See also AAR71482. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 238 AA;

Query Match 100.0%; Score 1207; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-115;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60

Qy         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120

Qy        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180

Qy        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
              |||||||||||||||||||||||||||||||||||||||
Db        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
```

RESULT 2
AAG79439
ID AAG79439 standard; protein; 238 AA.
XX
AC AAG79439;
XX
DT 25-OCT-2002 (first entry)
XX
DE Breast/colon cancer associated protein.

XX
KW Breast; colon; cancer; diagnosis; colorectal; prognosis; gene therapy;
KW CHA4; CBK8.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 6. .238
PT /note= "Encoded by ABA00091"
XX
PN WO200259609-A2.
XX
PD 01-AUG-2002.
XX
PF 10-DEC-2001; 2001WO-US048368.
XX
PR 08-DEC-2000; 2000US-00733756.
PR 08-DEC-2000; 2000US-00733757.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Gish KC, Wilson KE;
XX
DR WPI; 2002-599815/64.
DR N-PSDB; ABA00091.
XX
PT Diagnosing cancer by determining the expression profile gene that encodes
PT the CHA4 or CBK8 protein, useful for the treatment and prognosis of
PT breast and colorectal cancer.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The sequences given in ABA00086-92 represent sequences which have been
CC associated with either breast and/or colon cancer. These sequences may be
CC used in the method of the invention for diagnosing cancer. The method
CC comprises determining the expression of one or more specified genes in a
CC colorectal tissue sample, or in a breast tissue sample, of a first
CC individual. The method of the invention is useful for the diagnosis,
CC treatment and prognosis of breast and/or colorectal cancer. The invention
CC also provides for methods useful for modulating activity and inhibiting
CC breast and/or colorectal cancer. Prior methods of diagnosing and
CC prognosticating breast and colorectal cancer have been limited and
CC problematic. The present methods of gene therapy using novel sequences
CC for the diagnosis and prognosis of cancer, including screening for drug
CC candidates and bioactive agents that modulates CHA4 or CBK8, makes their
CC use and purpose limitless and more specific especially in the area of
CC grading prognostic factors. This sequence represents CHA4 which has the
XX bioactivity of a breast cancer modulating protein
SQ Sequence 238 AA;

Query Match 100.0%; Score 1207; DB 5; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-115;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60
Db          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60

Qy         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120
Db         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120

Qy        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180
Db        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180

Qy        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
Db        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
```

RESULT 3

ABP97190
ID ABP97190 standard; protein; 238 AA.
XX
AC ABP97190;
XX
DT 01-JUL-2003 (first entry)
XX
DE Tumour-associated antigenic target protein TAT178 SEQ ID NO:72.
XX
KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;
KW cancer.
XX
OS Homo sapiens.
XX
PN WO2003024392-A2.
XX
PD 27-MAR-2003.
XX
PF 11-SEP-2002; 2002WO-US028859.
XX
PR 18-SEP-2001; 2001US-0323268P.
PR 19-OCT-2001; 2001US-0339227P.
PR 07-NOV-2001; 2001US-0336827P.
PR 20-NOV-2001; 2001US-0331906P.
PR 02-JAN-2002; 2002US-0345444P.
PR 03-APR-2002; 2002US-0369724P.
PR 19-AUG-2002; 2002US-0404809P.
XX
PA (GETH) GENENTECH INC.
XX
PI Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
PI Williams PM, Wu TD, Zhang Z;
XX
DR WPI; 2003-354551/33.
DR N-PSDB; ACC49508.
XX
PT New antibodies against tumor-associated antigenic target polypeptide,
PT useful for treating or diagnosing tumors or cancers in mammals, e.g.
PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell
PT carcinomas.
XX
PS Claim 2; Fig 72; 285pp; English.
XX
CC ACC49493 to ACC49552 encode the human tumour-associated antigenic target
CC (TAT) proteins given in ABP97175 to ABP97234. The present invention
CC describes an isolated antibody that binds to a polypeptide having at
CC least 80 % sequence identity to any of the 60 150-800 residue amino acid
CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking
CC its associated signal peptide, encoded by any of the 60 2000-3000 base
CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have
CC cytostatic activity. The antibody can be used for treating or diagnosing
CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast
CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal
XX cell carcinomas, or thyroid cancer
SQ Sequence 238 AA;

Query Match 100.0%; Score 1207; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-115;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60
Db          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60

Qy         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120
Db         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120

Qy        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180
Db        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180
```

```
Qy        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
              |||||||||||||||||||||||||||||||||||||||||
Db        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
```

RESULT 4
ADN38812
ID ADN38812 standard; protein; 238 AA.
XX
AC ADN38812;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:130.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR N-PSDB; ADN38811.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO 130; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularistaion syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 238 AA;

```
  Query Match             100.0%;  Score 1207;  DB 7;  Length 238;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-115;
  Matches  219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60

Qy         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120

Qy        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180

Qy        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
              |||||||||||||||||||||||||||||||||||||||||
Db        181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
```

RESULT 5
ADQ21666
ID ADQ21666 standard; protein; 238 AA.
XX
AC ADQ21666;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4486.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4486; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the

CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 238 AA;

Query Match 100.0%; Score 1207; DB 8; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-115;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60
Db 1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60

Qy 61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120
Db 61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120

Qy 121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180
Db 121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180

Qy 181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
Db 181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219

RESULT 6
AAR82605
ID AAR82605 standard; protein; 234 AA.
XX
AC AAR82605;
XX
DT 16-MAY-1996 (first entry)
XX
DE Eph transmembrane tyrosine kinase family ligand, Efl-2.
XX
KW Efl-2; EHK1-L; Eph transmembrane tyrosine kinase family ligand;
KW neurological disorder; identification; diagnosis.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide Location/Qualifiers
FT 1. .30
FT /label= signal_peptide
FT Misc-difference 158
FT Region /note= "residue borders main conserved regions"
FT 218. .235
FT /note= "carboxy terminal hydrophobic GPI-recognition
XX tail"
PN WO9527060-A2.
XX
PD 12-OCT-1995.
XX
PF 04-APR-1995; 95WO-US004208.
XX
PR 04-APR-1994; 94US-00222075.
PR 12-APR-1994; 94US-00229402.
PR 01-SEP-1994; 94US-00299567.
PR 21-OCT-1994; 94US-00327423.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Davis S, Gale N, Aldrich TH, Maisonpierre PC, Goldfarb M;
PI Yancopoulos GD;
XX
DR WPI; 1995-358635/46.
DR N-PSDB; AAT03883.
XX
PT Ligands which bind Eph family receptors - used in the diagnosis of
PT neurological disorders.
XX
PS Disclosure; Fig 2; 58pp; English.
XX
CC Efl-2 (also known as EHK-1L) is an Eph transmembrane tyrosine kinase
CC family ligand. It has homology with B61 (Efl-1) (see AAR82604). Efl-2
CC ends in a C-terminal hydrophobic sequence that appears to to be a
CC recognition sequence allowing it to be GPI-linked and thus lacking in an
CC intracellular domain. Efl-2 is useful for identifying other ligands for
CC Ehk-1, -2, -3, Eck and Elk receptors. The ligands are useful in promoting
CC a differential function and/or influencing the phenotye, such as growth
CC and/or proliferation, of receptor bearing cells. They may be used in the
CC diagnosis, and treatment of neurological disorders
XX
SQ Sequence 234 AA;

Query Match 97.1%; Score 1172; DB 2; Length 234;
Best Local Similarity 98.2%; Pred. No. 7.2e-112;
Matches 215; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60
Db 1 MAAAPLLLLLLLVPVPLLPLLAQGPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60

Qy 61 IYCPHYNSSGVGPGAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120
Db 61 IYCPHYNSS----GAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 116

Qy 121 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 180
Db 117 KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT 176

Qy 181 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 219
Db 177 LPQFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKR 215

Search completed: April 19, 2005, 23:13:33
Job time : 59.9368 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 09:30:04 ; Search time 391.37 Seconds
(without alignments)
8258.620 Million cell updates/sec

Title: US-09-904-954-3_COPY_28_573
Perfect score: 546
Sequence: 1 ATGCGGCTGCTGCCCCTGCT.....GAGGGGGGGACACTCCCAGC 546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	546	100.0	636	2	AAQ85888	Aaq85888 Human hek
2	546	100.0	1181	8	ABZ34863	Abz34863 Coding se
3	544.4	99.7	606	6	ABV78135	Abv78135 Human eph
4	544.4	99.7	606	6	ABZ35711	Abz35711 Human eph
5	544.4	99.7	606	6	ABX09954	Abx09954 Human eph
6	544.4	99.7	606	6	ABL91676	Abl91676 Human pol

ALIGNMENTS

RESULT 1
AAQ85888
ID AAQ85888 standard; cDNA to mRNA; 636 BP.
XX
AC AAQ85888;
XX
DT 25-MAR-2003 (revised)
DT 03-OCT-1995 (first entry)
XX
DE Human hek-L protein cDNA clone C6.
XX
KW Ligand; cell surface; tyrosine kinase receptor; tumorigenesis; immunogen;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 28. .633
FT /*tag= a
FT sig_peptide 28. .93
FT /*tag= b
FT mat_peptide 94. .630
FT /*tag= c
XX
PN WO9506065-A1.
XX
PD 02-MAR-1995.
XX
PF 17-AUG-1994; 94WO-US009282.
XX
PR 20-AUG-1993; 93US-00109745.
PR 30-AUG-1993; 93US-00114426.
PR 03-DEC-1993; 93US-00161132.
PR 09-MAY-1994; 94US-00240124.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Beckmann MP, Cerretti DP;
XX
DR WPI; 1995-106811/14.
DR P-PSDB; AAR71482.
XX
PT New isolated DNA encoding hek-L protein or its fusion products - useful
PT as assay reagent or for carrying therapeutic and diagnostic compounds to
PT leukaemia cells.
XX
PS Claim 3; Page 37; 45pp; English.
XX
CC The sequence is that of a clone encoding hek-L protein, a protein that
CC can bind hek (a cell surface receptor tyrosine kinase). Hek-L is the
CC first known ligand for hek and can be used to study cellular processes
CC regulated by hek (which may be involved in tumorigenesis). It is also an
CC immunogen for antibody production, as a reagent for detecting hek or hek-
CC L in in vitro assays, to determine binding of hek proteins, to purify hek
CC proteins, and to carry diagnostic or cytotoxic agents to particular
CC leukaemia cells that express the hek antigen. Hek-L also binds the elk
CC tyrosine kinase receptors. See also AAQ85887. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 636 BP; 102 A; 202 C; 186 G; 146 T; 0 U; 0 Other;

Query Match 100.0%; Score 546; DB 2; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.1e-138;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 87

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 147

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 207

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 267

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
```

```
Db        268 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 327
Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
Db        328 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
Db        388 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
Db        448 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
Db        508 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
Qy        541 CCCAGC 546
Db        568 CCCAGC 573
```

RESULT 2
ABZ34863
ID ABZ34863 standard; cDNA; 1181 BP.
XX
AC ABZ34863;
XX
DT 04-FEB-2003 (first entry)
XX
DE Coding sequence SEQ ID 221, differentially expressed in osteogenesis.
XX
KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
KW osteoporosis; bone disease; upregulator; human; ephin-ephrin; ss.
XX
OS Homo sapiens.
XX
PN WO200281745-A2.
XX
PD 17-OCT-2002.
XX
PF 05-APR-2002; 2002WO-IB002211.
XX
PR 05-APR-2001; 2001US-0281400P.
XX
PA (AVET) AVENTIS PHARMA SA.
XX
PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
PI Connolly T, Jackson A, Bushnell SE, Rawadi G;
XX
DR WPI; 2003-058567/05.
XX
PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
PT useful for bone disease therapy in subject.
XX
PS Claim 1; Page 214; 237pp; English.
XX
CC The present invention relates to novel nucleotide sequences, which are
CC differentially expressed in models of osteogenesis upon being put in
CC contact with a stimulator of osteogenesis. The present sequence is one
CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
CC disease in a patient, promoting osteogenesis and/or preventing
CC osteoporosis/bone disease. The present sequence encodes a Ephin-ephrin
CC family protein
XX
SQ Sequence 1181 BP; 234 A; 353 C; 331 G; 263 T; 0 U; 0 Other;

Query Match   100.0%; Score 546; DB 8; Length 1181;
Best Local Similarity 100.0%; Pred. No. 2.5e-138;
Matches 546; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
Db         28 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
Db         88 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
Db        148 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
Db        208 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
Db        268 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
Db        328 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
Db        388 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
Db        448 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
Db        508 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
Qy        541 CCCAGC 546
Db        568 CCCAGC 573
```

RESULT 3
ABV78135
ID ABV78135 standard; DNA; 606 BP.
XX
AC ABV78135;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human ephrin-A3 DNA SEQ ID NO 19.
XX
KW RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic;
KW virucide; protozoacide; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200255693-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-EP000152.
XX
PR 09-JAN-2001; 2001DE-01000586.
PR 26-OCT-2001; 2001DE-01055280.
PR 29-NOV-2001; 2001DE-01058411.
PR 07-DEC-2001; 2001DE-01060151.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-590671/63.
XX

PT Inhibiting expression of target gene, useful e.g. for inhibiting
PT oncogenes, by administering double-stranded RNA complementary to the
PT target and having an overhang.
XX
PS Claim 10; Page 123; 203pp; German.
XX
CC The invention relates to inhibiting expression of a target gene (I) in a
CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
CC structure of at most 49 consecutive bases. At least part of one strand
CC (as1) of dsRNA1 is complementary to (I) and at least one end of dsRNA1
CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC in humans, also genes in Plasmodium or in viruses or viroids that are
CC pathogenic for humans, animals or plants. Introducing an overhang into
CC dsRNA greatly increases effectiveness for inhibiting gene expression,
CC both in vivo and in vitro and also increases stability and thus the
CC effective concentration inside the cell. The present sequence is that of
CC a gene related to the invention
XX
SQ Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

Query Match 99.7%; Score 544.4; DB 6; Length 606;
Best Local Similarity 99.8%; Pred. No. 5.5e-138;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 360

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Qy        541 CCCAGC 546
              ||||||
Db        541 CCCAGC 546
```

RESULT 4
ABZ35711
ID ABZ35711 standard; DNA; 606 BP.
XX
AC ABZ35711;
XX
DT 07-FEB-2003 (first entry)
XX
DE Human ephrin A3 encoding polynucleotide SEQ ID NO 19.
XX
KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW Hepatitis C virus; human papilloma virus; gene; ds.
XX
OS Homo sapiens.
XX
PN DE10100588-A1.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2001; 2001DE-01000588.
XX
PR 09-JAN-2001; 2001DE-01000588.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-683450/74.
XX
PT Inhibiting expression of target genes, useful e.g. for treating tumors,
PT by introducing into cells two double-stranded RNAs that are complementary
PT to the target.
XX
PS Claim 13; Page 19-20; 100pp; German.
XX
CC The invention relates to inhibiting expression of a target gene in a cell
CC by introducing at least two oligoribonucleotides (dsRNAI and II), both
CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
CC pairs. At least part of one strand (S1, S2) of the ds structures in each
CC of dsRNAI and II are complementary to regions in the target gene. The
CC method uses antisense inhibition of gene expression using double stranded
CC RNA inhibition (RNAi). The method is particularly used to treat tumours
CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
CC humans, animals or plants). The method provides more effective inhibition
CC of expression than known methods using a single dsRNA, even at very low
CC concentrations. When dsRNA has at least one unpaired nucleotide at the
CC end, stability (and thus effective concentration in the cell) is improved
CC and efficiency can be increased further by pretreating the cells with
CC interferon. The present sequence is that of a target DNA of the invention
XX
SQ Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

Query Match 99.7%; Score 544.4; DB 6; Length 606;
Best Local Similarity 99.8%; Pred. No. 5.5e-138;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
```

```
Db        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 360

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Qy        541 CCCAGC 546
              ||||||
Db        541 CCCAGC 546
```

RESULT 5
ABX09954
ID ABX09954 standard; DNA; 606 BP.
XX
AC ABX09954;
XX
DT 23-JAN-2003 (first entry)
XX
DE Human ephrin A3 DNA fragment SEQ ID 19.
XX
KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW prion; inhibition; human; ds.
XX
OS Homo sapiens.
XX
PN DE10100587-C1.
XX
PD 21-NOV-2002.
XX
PF 09-JAN-2001; 2001DE-01000587.
XX
PR 09-JAN-2001; 2001DE-01000587.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-742209/81.
XX
PT Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT introduction of complementary double-stranded oligoribonucleotide, after
PT treating the cell with interferon.
XX
PS Disclosure; Page 24-25; 98pp; German.
XX
CC This invention describes a novel method for inhibiting expression of a
CC target gene by introducing into the cell that contains the target gene at
CC least one oligoribonucleotide (dsRNAI) that has a double-stranded (ds)
CC structure of not more than 49 consecutive nucleotides (nt), where at
CC least a segment of one strand of the ds structure is complementary with
CC the target gene and the cells are treated with interferon before
CC introduction of dsRNAI. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX09936-ABX10075 represent
CC gene fragments used to illustrate the method of the invention
XX
SQ Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

Query Match 99.7%; Score 544.4; DB 6; Length 606;

Best Local Similarity 99.8%; Pred. No. 5.5e-138;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 360

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Qy        541 CCCAGC 546
              ||||||
Db        541 CCCAGC 546
```

RESULT 6
ABL91676
ID ABL91676 standard; DNA; 606 BP.
XX
AC ABL91676;
XX
DT 28-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 19.
XX
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS Homo sapiens.
XX
PN DE10100586-C1.
XX
PD 11-APR-2002.
XX
PF 09-JAN-2001; 2001DE-01000586.
XX
PR 09-JAN-2001; 2001DE-01000586.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-270454/32.

XX
PT Inhibiting gene expression in cells, useful for e.g. treating tumors, by
PT introducing double-stranded complementary oligoRNA having unpaired
PT terminal bases.
XX
PS Claim 13; Page 20-21; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a single-
CC stranded segment of 1-4 nt. The method provides oligoribonucleotides for
CC antisense inhibition of gene expression useful e.g. for treating tumours
CC but the oligoribonucleotides may also be directed against genes present
CC in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans,
CC animals or plants) or against cytokine, Id, developmental or prion genes.
CC The method provides more effective inhibition of gene expression than use
CC of known oligonucleotides, probably because the unpaired overhang
CC increases stability and thus intracellular concentration
XX
SQ Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

Query Match 99.7%; Score 544.4; DB 6; Length 606;
Best Local Similarity 99.8%; Pred. No. 5.5e-138;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 360

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Qy        541 CCCAGC 546
              ||||||
Db        541 CCCAGC 546
```

Search completed: April 19, 2005, 16:26:10
Job time : 391.513 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 09:30:33 ; Search time 118.434 Seconds
(without alignments)
7543.483 Million cell updates/sec

Title: US-09-904-954-3_COPY_28_573
Perfect score: 546
Sequence: 1 ATGCGGCTGCTGCCCCTGCT.....GAGGGGGGGACACTCCCAGC 546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	546	100.0	636	1	US-08-240-124-3	Sequence 3, Appli
2	546	100.0	636	1	US-08-453-943-3	Sequence 3, Appli
3	546	100.0	636	2	US-09-057-121-3	Sequence 3, Appli
4	546	100.0	636	3	US-09-358-734-3	Sequence 3, Appli
5	544.4	99.7	1182	4	US-09-949-016-1291	Sequence 1291, Ap

ALIGNMENTS

RESULT 1
US-08-240-124-3
; Sequence 3, Application US/08240124
; Patent No. 5516658
; GENERAL INFORMATION:
; APPLICANT: BECKMANN, M. P.
; APPLICANT: CERRETTI, DOUGLAS P.
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
; TITLE OF INVENTION: RECEPTOR HEK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,124
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,132
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/114,426
; FILING DATE: 30-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/109,745
; FILING DATE: 20-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SEESE, KATHRYN A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2814-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: hek-L C6
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 94..630
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..633
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 28..93
US-08-240-124-3

Query Match 100.0%; Score 546; DB 1; Length 636;
Best Local Similarity 100.0%; Pred. No. 1.9e-142;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 87

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 147

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 207

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 267

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 327

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
```

```
Db        328 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 387

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420

Db        388 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Db        448 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Db        508 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Qy        541 CCCAGC 546

Db        568 CCCAGC 573
```

```
RESULT 2
US-08-453-943-3
; Sequence 3, Application US/08453943
; Patent No. 5738844
;  GENERAL INFORMATION:
;    APPLICANT:  BECKMANN, M. P.
;    APPLICANT:  CERRETTI, DOUGLAS P.
;    TITLE OF INVENTION:  CYTOKINE THAT BINDS THE CELL SURFACE
;    TITLE OF INVENTION:  RECEPTOR HEK
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  IMMUNEX CORPORATION
;      STREET:  51 UNIVERSITY STREET
;      CITY:  SEATTLE
;      STATE:  WASHINGTON
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Apple System 7.1
;      SOFTWARE:  Microsoft Word for Apple, Version 5.1a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/453,943
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/240,124
;      FILING DATE:  09-MAY-1994
;      APPLICATION NUMBER:  US 08/161,132
;      FILING DATE:  03-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/114,426
;      FILING DATE:  30-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/109,745
;      FILING DATE:  20-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SEESE, KATHRYN A.
;      REGISTRATION NUMBER:  32,172
;      REFERENCE/DOCKET NUMBER:  2814-C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 587-0430
;      TELEFAX:  (206) 233-0644
;      TELEX:  756822
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  636 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  hek-L C6
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  94..630
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  28..633
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  28..93
US-08-453-943-3

  Query Match             100.0%;  Score 546;  DB 1;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-142;
  Matches  546;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60

Db         28 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120

Db         88 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 147

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180

Db        148 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 207

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240

Db        208 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 267

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300

Db        268 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 327

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360

Db        328 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 387

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420

Db        388 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 447

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Db        448 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 507

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Db        508 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 567

Qy        541 CCCAGC 546

Db        568 CCCAGC 573
```

```
RESULT 3
US-09-057-121-3
; Sequence 3, Application US/09057121
; Patent No. 5969110
;  GENERAL INFORMATION:
;    APPLICANT:  BECKMANN, M. P.
;    APPLICANT:  CERRETTI, DOUGLAS P.
;    TITLE OF INVENTION:  CYTOKINE THAT BINDS THE CELL SURFACE
;    TITLE OF INVENTION:  RECEPTOR HEK
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  IMMUNEX CORPORATION
;      STREET:  51 UNIVERSITY STREET
;      CITY:  SEATTLE
```

```
;       STATE:  WASHINGTON
;       COUNTRY:  USA
;       ZIP:  98101
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  Apple Macintosh
;       OPERATING SYSTEM:  Apple System 7.1
;       SOFTWARE:  Microsoft Word for Apple, Version 5.1a
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/057,121
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/240,124
;       FILING DATE:
;       APPLICATION NUMBER:  US 08/161,132
;       FILING DATE:  03-DEC-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/114,426
;       FILING DATE:  30-AUG-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/109,745
;       FILING DATE:  20-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  SEESE, KATHRYN A.
;       REGISTRATION NUMBER:  32,172
;       REFERENCE/DOCKET NUMBER:  2814-C
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 587-0430
;       TELEFAX:  (206) 233-0644
;       TELEX:  756822
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  636 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     IMMEDIATE SOURCE:
;       CLONE:  hek-L C6
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  94..630
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  28..633
;     FEATURE:
;       NAME/KEY:  sig_peptide
;       LOCATION:  28..93
US-09-057-121-3

  Query Match             100.0%;  Score 546;  DB 2;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-142;
  Matches  546;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 87

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 147

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 207

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 267

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 327

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 387

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 447

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 507

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 567

Qy        541 CCCAGC 546
              ||||||
Db        568 CCCAGC 573


RESULT 4
US-09-358-734-3
; Sequence 3, Application US/09358734
; Patent No. 6274117
;  GENERAL INFORMATION:
;    APPLICANT:  BECKMANN, M. P.
;    APPLICANT:  CERRETTI, DOUGLAS P.
;    TITLE OF INVENTION:  CYTOKINE THAT BINDS THE CELL SURFACE
;    TITLE OF INVENTION:  RECEPTOR HEK
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  IMMUNEX CORPORATION
;      STREET:  51 UNIVERSITY STREET
;      CITY:  SEATTLE
;      STATE:  WASHINGTON
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Apple System 7.1
;      SOFTWARE:  Microsoft Word for Apple, Version 5.1a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/358,734
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/240,124
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/114,426
;      FILING DATE:  30-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/109,745
;      FILING DATE:  20-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SEESE, KATHRYN A.
;      REGISTRATION NUMBER:  32,172
;      REFERENCE/DOCKET NUMBER:  2814-C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 587-0430
;      TELEFAX:  (206) 233-0644
;      TELEX:  756822
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  636 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;     TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     IMMEDIATE SOURCE:
;       CLONE:  hek-L C6
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  94..630
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  28..633
;     FEATURE:
;       NAME/KEY:  sig_peptide
;       LOCATION:  28..93
US-09-358-734-3

  Query Match                 100.0%;  Score 546;  DB 3;  Length 636;
  Best Local Similarity       100.0%;  Pred. No. 1.9e-142;
  Matches  546;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 87

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 147

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 207

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 267

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 327

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 387

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 447

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 507

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 567

Qy        541 CCCAGC 546
              ||||||
Db        568 CCCAGC 573


RESULT 5
US-09-949-016-1291
, Sequence 1291, Application US/09949016
, Patent No. 6812339
, GENERAL INFORMATION:
,  APPLICANT: VENTER, J. Craig et al.
,  TITLE OF INVENTION:  POLYMORPHISMS IN KNOWN GENES ASSOCIATED
,  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
,  FILE REFERENCE: CL001307
,  CURRENT APPLICATION NUMBER: US/09/949,016
,  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1291
;  LENGTH: 1182
;  TYPE: DNA
;  ORGANISM: Human
US-09-949-016-1291

  Query Match                 99.7%;  Score 544.4;  DB 4;  Length 1182;
  Best Local Similarity       99.8%;  Pred. No. 6.6e-142;
  Matches  545;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 87

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 147

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 207

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 267

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 327

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 387

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 447

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 507

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 567

Qy        541 CCCAGC 546
              ||||||
Db        568 CCCAGC 573


Search completed: April 19, 2005, 22:12:42
Job time : 118.434 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 09:30:04 ; Search time 2539.69 Seconds
(without alignments)
8183.329 Million cell updates/sec

Title: US-09-904-954-3_COPY_28_573
Perfect score: 546
Sequence: 1 ATGCGGCTGCTGCCCCTGCT.....GAGGGGGGGACACTCCCAGC 546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 10

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	544.4	99.7	619	7	CN289865	CN289865 170006002
2	544.4	99.7	636	7	CN289863	CN289863 170006000
3	544.4	99.7	965	5	BQ928753	BQ928753 AGENCOURT
4	537	98.4	898	6	CA454936	CA454936 AGENCOURT
5	532.4	97.5	707	2	BE883793	BE883793 601506189
6	532.4	97.5	912	4	BI868252	BI868252 603392205
7	521.4	95.5	838	2	BE890843	BE890843 601431292
8	499.4	91.5	663	7	CN289864	CN289864 170004247
9	498.2	91.2	834	6	CB996846	CB996846 AGENCOURT
10	497.4	91.1	592	6	CD671399	CD671399 fg04c02.y

ALIGNMENTS

RESULT 1
CN289865
LOCUS CN289865 619 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000600263646 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN289865
VERSION CN289865.1 GI:47306279
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 619)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 619 Std Error: 0.00.
FEATURES Location/Qualifiers
source 1. .619
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"
/clone_lib="GRN_PREHEP"
/note="oligo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"
ORIGIN

Query Match 99.7%; Score 544.4; DB 7; Length 619;
Best Local Similarity 99.8%; Pred. No. 1.3e-129;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 70

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 130

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 190

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 250

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 310

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        311 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 370

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 430

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 490

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 550

Qy        541 CCCAGC 546
              ||||||
Db        551 CCCAGC 556
```

RESULT 2
CN289863
LOCUS CN289863 636 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000600021389 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN289863
VERSION CN289863.1 GI:47306277
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 636)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 636 Std Error: 0.00.
FEATURES Location/Qualifiers
source 1. .636
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"
/clone_lib="GRN_PREHEP"
/note="oligo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"
ORIGIN

Query Match 99.7%; Score 544.4; DB 7; Length 636;
Best Local Similarity 99.8%; Pred. No. 1.3e-129;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 102

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 162

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCCAC 222

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 282

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 342

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 402

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 462
```

Qy 421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
Db 463 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480

Qy 481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
Db 523 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540

Qy 541 CCCAGC 546
Db 583 CCCAGC 588

RESULT 3
BQ928753
LOCUS BQ928753 965 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT_10036172 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481391 5', mRNA sequence.
ACCESSION BQ928753
VERSION BQ928753.1 GI:22343784
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 965)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2662 row: k column: 24
High quality sequence stop: 635.
FEATURES Location/Qualifiers
source 1. .965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481391"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 99.7%; Score 544.4; DB 5; Length 965;
Best Local Similarity 99.8%; Pred. No. 1.4e-129;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 115

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 175

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        176 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 235

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 295

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 355

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        356 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 415

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 475

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 535

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 595

Qy        541 CCCAGC 546
              ||||||
Db        596 CCCAGC 601
```

RESULT 4
CA454936
LOCUS CA454936 898 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10714679 MAPcL Homo sapiens cDNA clone IMAGE:6722424 5', mRNA sequence.
ACCESSION CA454936
VERSION CA454936.1 GI:24905152
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 898)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Kristi A. Egland, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14285 row: b column: 24
High quality sequence stop: 473.

FEATURES Location/Qualifiers
source 1. .898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722424"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Egland, James J. Vincent, Robert Strausberg, Bungkook Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Query Match 98.4%; Score 537; DB 6; Length 898;
Best Local Similarity 98.9%; Pred. No. 1.1e-127;
Matches 540; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 115

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 175

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 235

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 295

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 355

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        356 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 415

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 475

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              |||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db        476 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCANGGAGAAGAAGTCTGAGTCA 535

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
              |||||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||
Db        536 GCCCATCCTGTTGGGAGCCCTGGAGAAAGTGGAACATCAGGGTGGCGAGGGGGGGACACT 595

Qy        541 CCCAGC 546
              ||||||
Db        596 TCCAGC 601
```

RESULT 5
BE883793
LOCUS BE883793 707 bp mRNA linear EST 20-OCT-2000
DEFINITION 601506189F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3907510 5', mRNA sequence.
ACCESSION BE883793
VERSION BE883793.1 GI:10332569
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 707)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9718 row: b column: 23
High quality sequence stop: 706.

FEATURES Location/Qualifiers
source 1. .707
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3907510"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. "

ORIGIN

Query Match 97.5%; Score 532.4; DB 2; Length 707;
Best Local Similarity 99.6%; Pred. No. 1.6e-126;
Matches 544; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy      1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 108

Qy     61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 168

Qy    121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 228

Qy    181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 TACGAAGGCCCA-GGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 287

Qy    241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 347

Qy    301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    348 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 407

Qy    361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    408 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 467

Qy    421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    468 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 527

Qy    481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    528 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 587

Qy    541 CCCAGC 546
          ||||||
Db    588 CCCAGC 593
```

RESULT 6
BI868252
LOCUS BI868252 912 bp mRNA linear EST 11-OCT-2001
DEFINITION 603392205F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402208 5', mRNA sequence.
ACCESSION BI868252
VERSION BI868252.1 GI:16041925
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 912)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12025 row: j column: 01
High quality sequence stop: 806.

FEATURES Location/Qualifiers
source 1. .912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5402208"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 97.5%; Score 532.4; DB 4; Length 912;
Best Local Similarity 99.6%; Pred. No. 1.7e-126;
Matches 544; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy      1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     39 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 98

Qy     61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 158

Qy    121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    159 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 218

Qy    181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    219 TACGAAGGCCCA-GGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 277

Qy    241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    278 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 337

Qy    301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    338 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 397

Qy    361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    398 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 457

Qy    421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    458 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 517

Qy    481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    518 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACT 577
```

```
Qy        541 CCCAGC 546
              ||||||
Db        578 CCCAGC 583


RESULT 7
BE890843
LOCUS       BE890843                838 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601431292F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3916487 5',
            mRNA sequence.
ACCESSION   BE890843
VERSION     BE890843.1  GI:10349570
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9741  row: h  column: 24
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .838
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3916487"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             95.5%;  Score 521.4;  DB 2;  Length 838;
  Best Local Similarity   99.5%;  Pred. No. 1.2e-123;
  Matches  544;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 130

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 190

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 250

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        251 TACGAAGGCCCA-GGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 309

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 369

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||

Db        370 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 429

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 489

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 549

Qy        481 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGA-GGGGGGGACAC 539
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        550 GCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGACGGGGGGGACAC 609

Qy        540 TCCCAGC 546
              |||||||
Db        610 TCCCAGC 616


RESULT 8
CN289864
LOCUS       CN289864                663 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424723908 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN289864
VERSION     CN289864.1  GI:47306278
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 663)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 663   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .663
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, cell lines H1, H7, and
                     H9"
                     /clone_lib="GRN_ES"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from undifferentiated hES cell lines H1 (p32), H7 (p29),
                     and H9 (p26) maintained in feeder-free conditions"
ORIGIN

  Query Match             91.5%;  Score 499.4;  DB 7;  Length 663;
  Best Local Similarity   99.8%;  Pred. No. 5.3e-118;
  Matches  500;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         46 CTCGGCTCCCCTCTGCGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGT 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 CTCGGCTCCCCTCTGCGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGT 71

Qy        106 AACCCCAGGTTGCTTCGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGAC 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 AACCCCAGGTTGCTTCGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGAC 131

Qy        166 ATTGTCTGCCCCCACTACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTG 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 ATTGTCTGCCCCCACTACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTG 191
```

```
Qy        226 TACATGGTGGACTGGCCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAG 285
Db        192 TACATGGTGGACTGGCCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAG 251

Qy        286 CGCTGGGTGTGCTCCCTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTC 345
Db        252 CGCTGGGTGTGCTCCCTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTC 311

Qy        346 ACACCTTTCTCCCTCGGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTG 405
Db        312 ACACCCTTCTCCCTCGGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTG 371

Qy        406 CCCACTCCAGAGAGTTCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAG 465
Db        372 CCCACTCCAGAGAGTTCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAG 431

Qy        466 AGGAAGTCTGAGTCAGCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGG 525
Db        432 AGGAAGTCTGAGTCAGCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGG 491

Qy        526 CGAGGGGGGGACACTCCCAGC 546
Db        492 CGAGGGGGGGACACTCCCAGC 512
```

RESULT 9
CB996846
LOCUS CB996846 834 bp mRNA linear EST 01-MAY-2003
DEFINITION AGENCOURT_13666148 NIH_MGC_148 Homo sapiens cDNA clone IMAGE:30332601 5', mRNA sequence.
ACCESSION CB996846
VERSION CB996846.1 GI:30291366
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM349 row: o column: 10
High quality sequence stop: 514.
FEATURES Location/Qualifiers
source 1. .834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30332601"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B TonA"
/clone_lib="NIH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site_1: alI-XhoI; Site_2: BamH; Library is oligo-dT primed and directionally cloned using primer
5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-lenght clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 91.2%; Score 498.2; DB 6; Length 834;
Best Local Similarity 97.2%; Pred. No. 1.1e-117;
Matches 528; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

```
Qy          1 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
Db         99 ATGCGGCTGCTGCCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 158

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
Db        159 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 218

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
Db        219 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 278

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
Db        279 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 338

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
Db        339 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 398

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
Db        399 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 458

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
Db        459 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 518

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCT-GCAAGGAGAGGAAGTCTGAGTC 479
Db        519 TCTGGCCAGTGCTTGAGGCTCCCGGTGTCTGTCTGCTGGCAGGGAGAGGAATTCTGAGTC 578

Qy        480 AGCCCATCCTGTTGGGAGCCCT-GGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACA 538
Db        579 ATCCCATCCTGGTGGGAGCCCTGGGAGAGAGTGGCACATCTAGGTGGCCAAGGGGGGGA 638

Qy        539 CTC 541
Db        639 CAC 641
```

RESULT 10
CD671399
LOCUS CD671399 592 bp mRNA linear EST 24-JUN-2003
DEFINITION fg04c02.y1 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone fg04c02 5', mRNA sequence.
ACCESSION CD671399
VERSION CD671399.1 GI:32173139
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 592)
AUTHORS Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A., Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of adult human iris for the NEIBank Project: steroid-response factors and similarities with retinal pigment epithelium
JOURNAL Mol. Vis. 8 (4), 185-195 (2002)
MEDLINE 22103462
PUBMED 12107412
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452

Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 04 row: c column: 02
Seq primer: M13RP1 reverse primer (ABI).

FEATURES Location/Qualifiers
source 1. .592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fg04c02"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Iris cDNA (Normalized): fg"
/note="Organ: Eye; Vector: pCMVSPORT6; A human iris library (bx) was normalized by self-subtraction. One portion of double stranded plasmid DNA representing the library was linearized by NotI. This NotI digested library was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Life Technologies). Single-stranded DNA (1 mg) was hybridized (C0t 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center(NISC)."

ORIGIN

Query Match 91.1%; Score 497.4; DB 6; Length 592;
Best Local Similarity 99.8%; Pred. No. 1.7e-117;
Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 ATGCGGCTGCTGCCCCTGCTGCGGACTGTCCTCTGGGCCGCGTTCCTCGGCTCCCCTCTG 153

Qy         61 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 CGCGGGGGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTT 213

Qy        121 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 CGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCAC 273

Qy        181 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 TACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGG 333

Qy        241 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCC 393

Qy        301 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        394 CTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTC 453

Qy        361 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 GGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGT 513

Qy        421 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCA 573

Qy        481 GCCCATCCTGTTGGGAGCC 499
              |||||||||||||||||||
Db        574 GCCCATCCTGTTGGGAGCC 592
```

Search completed: April 19, 2005, 21:57:12
Job time : 2540.83 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 19, 2005, 23:10:12 ; Search time 35.3289 Seconds
(without alignments)
1684.011 Million cell updates/sec

Title: US-09-904-954-4_COPY_1_179
Perfect score: 989
Sequence: 1 MRLLPLLRTVLWAAFLGSPL.....SAHPVGSPGESGTSGWRGGD 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	989	100.0	201	9	US-09-904-954-4	Sequence 4, Appli
2	989	100.0	201	17	US-10-698-907-12	Sequence 12, Appl
3	971	98.2	201	13	US-10-138-787-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-904-954-4
; Sequence 4, Application US/09904954
; Patent No. US20020010325A1
; GENERAL INFORMATION:
; APPLICANT: BECKMANN, M. P.
; CERRETTI, DOUGLAS P.
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
; RECEPTOR HEK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/904,954
; FILING DATE: 12-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/240,124
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/114,426
; FILING DATE: 30-AUG-1993
; APPLICATION NUMBER: US 08/109,745
; FILING DATE: 20-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SEESE, KATHRYN A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2814-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-904-954-4

Query Match 100.0%; Score 989; DB 9; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.4e-96;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
```

RESULT 2
US-10-698-907-12
; Sequence 12, Application US/10698907
; Publication No. US20050049194A1
; GENERAL INFORMATION:
; APPLICANT: Frisen, Jonas
; APPLICANT: Holmberg, Johan
; TITLE OF INVENTION: Use of Ephrins and Related Molecules to Regulate Cellular
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: 21882-529 UTIL
; CURRENT APPLICATION NUMBER: US/10/698,907

; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: US 60/460,488
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 10/291,290
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 60/393,272
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/345,206
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-698-907-12

Query Match 100.0%; Score 989; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.4e-96;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
```

RESULT 3
US-10-138-787-8
; Sequence 8, Application US/10138787
; Publication No. US20020172984A1
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Pawson, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASES
; FILE REFERENCE: 11757.23USWO
; CURRENT APPLICATION NUMBER: US/10/138,787
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US/09/214,631
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: PCT/CA97/00473
; PRIOR FILING DATE: 1997-07-04
; PRIOR APPLICATION NUMBER: 60/021,272
; PRIOR FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-787-8

Query Match 98.2%; Score 971; DB 13; Length 201;
Best Local Similarity 98.3%; Pred. No. 1.9e-94;
Matches 176; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||| |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 MRLLPLLITVLWAAFLGSPLRGGSSLRHVVYWNSSNPRSLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRGYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
```

Qy 121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
Db 121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179

Search completed: April 19, 2005, 23:22:25
Job time : 35.3289 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2005, 22:12:51 ; Search time 14.1316 Seconds
(without alignments)
945.555 Million cell updates/sec

Title: US-09-904-954-4_COPY_1_179
Perfect score: 989
Sequence: 1 MRLLPLLRTVLWAAFLGSPL.....SAHPVGSPGESGTSGWRGGD 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	989	100.0	201	1	US-08-240-124-4	Sequence 4, Appli
2	989	100.0	201	1	US-08-453-943-4	Sequence 4, Appli
3	989	100.0	201	2	US-09-057-121-4	Sequence 4, Appli
4	989	100.0	201	3	US-09-358-734-4	Sequence 4, Appli
5	989	100.0	210	4	US-09-949-016-7162	Sequence 7162, Ap
6	971	98.2	201	4	US-09-214-631-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-240-124-4
; Sequence 4, Application US/08240124
; Patent No. 5516658
; GENERAL INFORMATION:
; APPLICANT: BECKMANN, M. P.
; APPLICANT: CERRETTI, DOUGLAS P.
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
; TITLE OF INVENTION: RECEPTOR HEK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,124
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,132
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/114,426
; FILING DATE: 30-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/109,745
; FILING DATE: 20-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SEESE, KATHRYN A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2814-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-240-124-4

Query Match 100.0%; Score 989; DB 1; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.6e-107;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
```

RESULT 2
US-08-453-943-4
; Sequence 4, Application US/08453943
; Patent No. 5738844
; GENERAL INFORMATION:
; APPLICANT: BECKMANN, M. P.
; APPLICANT: CERRETTI, DOUGLAS P.
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
; TITLE OF INVENTION: RECEPTOR HEK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1

```
;       SOFTWARE:  Microsoft Word for Apple, Version 5.1a
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/453,943
;       FILING DATE:  30-MAY-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/240,124
;       FILING DATE:  09-MAY-1994
;       APPLICATION NUMBER:  US 08/161,132
;       FILING DATE:  03-DEC-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/114,426
;       FILING DATE:  30-AUG-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/109,745
;       FILING DATE:  20-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  SEESE, KATHRYN A.
;       REGISTRATION NUMBER:  32,172
;       REFERENCE/DOCKET NUMBER:  2814-C
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 587-0430
;       TELEFAX:  (206) 233-0644
;       TELEX:  756822
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  201 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-453-943-4

  Query Match                 100.0%;  Score 989;  DB 1;  Length 201;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-107;
  Matches  179;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179


RESULT 3
US-09-057-121-4
; Sequence 4, Application US/09057121
; Patent No. 5969110
; GENERAL INFORMATION:
;   APPLICANT:  BECKMANN, M. P.
;   APPLICANT:  CERRETTI, DOUGLAS P.
;   TITLE OF INVENTION:  CYTOKINE THAT BINDS THE CELL SURFACE
;   TITLE OF INVENTION:  RECEPTOR HEK
;   NUMBER OF SEQUENCES:  4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  IMMUNEX CORPORATION
;     STREET:  51 UNIVERSITY STREET
;     CITY:  SEATTLE
;     STATE:  WASHINGTON
;     COUNTRY:  USA
;     ZIP:  98101
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  Apple Macintosh
;     OPERATING SYSTEM:  Apple System 7.1
;     SOFTWARE:  Microsoft Word for Apple, Version 5.1a
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/057,121
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/240,124
;     FILING DATE:
;     APPLICATION NUMBER:  US 08/161,132
;     FILING DATE:  03-DEC-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/114,426
;     FILING DATE:  30-AUG-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/109,745
;     FILING DATE:  20-AUG-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  SEESE, KATHRYN A.
;     REGISTRATION NUMBER:  32,172
;     REFERENCE/DOCKET NUMBER:  2814-C
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (206) 587-0430
;     TELEFAX:  (206) 233-0644
;     TELEX:  756822
; INFORMATION FOR SEQ ID NO:  4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  201 amino acids
;     TYPE:  amino acid
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-09-057-121-4

  Query Match                 100.0%;  Score 989;  DB 2;  Length 201;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-107;
  Matches  179;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179


RESULT 4
US-09-358-734-4
; Sequence 4, Application US/09358734
; Patent No. 6274117
; GENERAL INFORMATION:
;   APPLICANT:  BECKMANN, M. P.
;   APPLICANT:  CERRETTI, DOUGLAS P.
;   TITLE OF INVENTION:  CYTOKINE THAT BINDS THE CELL SURFACE
;   TITLE OF INVENTION:  RECEPTOR HEK
;   NUMBER OF SEQUENCES:  4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  IMMUNEX CORPORATION
;     STREET:  51 UNIVERSITY STREET
;     CITY:  SEATTLE
;     STATE:  WASHINGTON
;     COUNTRY:  USA
;     ZIP:  98101
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  Apple Macintosh
;     OPERATING SYSTEM:  Apple System 7.1
;     SOFTWARE:  Microsoft Word for Apple, Version 5.1a
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/358,734
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  08/240,124
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/114,426
;      FILING DATE:  30-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/109,745
;      FILING DATE:  20-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SEESE, KATHRYN A.
;      REGISTRATION NUMBER:  32,172
;      REFERENCE/DOCKET NUMBER:  2814-C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 587-0430
;      TELEFAX:  (206) 233-0644
;      TELEX:  756822
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  201 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-358-734-4

  Query Match           100.0%;  Score 989;  DB 3;  Length 201;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-107;
  Matches  179;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179


RESULT 5
US-09-949-016-7162
; Sequence 7162, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7162
;   LENGTH: 210
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-7162

  Query Match           100.0%;  Score 989;  DB 4;  Length 210;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-107;
  Matches  179;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 69

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 129

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 188


RESULT 6
US-09-214-631-8
; Sequence 8, Application US/09214631
; Patent No. 6413730
; GENERAL INFORMATION:
;  APPLICANT: Holland, Sacha
;  APPLICANT: Mbamalu, Geraldine
;  APPLICANT: Pawson, Tony
;  TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
;  TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
;  TITLE OF INVENTION: TYROSINE KINASES
;  FILE REFERENCE: 11757.23USWO
;  CURRENT APPLICATION NUMBER: US/09/214,631
;  CURRENT FILING DATE: 1999-03-12
;  EARLIER APPLICATION NUMBER: PCT/CA97/00473
;  EARLIER FILING DATE: 1997-07-04
;  EARLIER APPLICATION NUMBER: 60/021,272
;  EARLIER FILING DATE: 1996-07-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 201
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-214-631-8

  Query Match           98.2%;  Score 971;  DB 4;  Length 201;
  Best Local Similarity  98.3%;  Pred. No. 3.3e-105;
  Matches  176;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||
Db          1 MRLLPLLITVLWAAFLGSPLRGGSSLRHVVYWNSSNPRSLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETFALYMVDWPGYESCQAEGPRGYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179


Search completed: April 19, 2005, 23:19:49
Job time : 14.1316 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 19, 2005, 15:35:11 ; Search time 48.9895 Seconds
(without alignments)
1413.163 Million cell updates/sec

Title: US-09-904-954-4_COPY_1_179
Perfect score: 989
Sequence: 1 MRLLPLLRTVLWAAFLGSPL.....SAHPVGSPGESGTSGWRGGD 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	989	100.0	201	2	AAR71482	Aar71482 Human hek

ALIGNMENTS

RESULT 1
AAR71482
ID AAR71482 standard; protein; 201 AA.
XX
AC AAR71482;
XX
DT 25-MAR-2003 (revised)
DT 03-OCT-1995 (first entry)
XX
DE Human hek-L protein.
XX
KW Ligand; cell surface; tyrosine kinase receptor; tumorigenesis; immunogen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .22
FT /note= "signal peptide"
FT Protein 23. .201
XX
PN WO9506065-A1.
XX
PD 02-MAR-1995.
XX
PF 17-AUG-1994; 94WO-US009282.
XX
PR 20-AUG-1993; 93US-00109745.
PR 30-AUG-1993; 93US-00114426.
PR 03-DEC-1993; 93US-00161132.
PR 09-MAY-1994; 94US-00240124.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Beckmann MP, Cerretti DP;
XX
DR WPI; 1995-106811/14.
DR N-PSDB; AAQ85888.
XX
PT New isolated DNA encoding hek-L protein or its fusion products - useful
PT as assay reagent or for carrying therapeutic and diagnostic compounds to
PT leukaemia cells.
XX
PS Claim 21; Page 38; 45pp; English.
XX
CC The sequence is that of a novel protein designated hek-L, a protein that
CC can bind hek (a cell surface receptor tyrosine kinase). Hek-L is the
CC first known ligand for hek and can be used to study cellular processes
CC regulated by hek (which may be involved in tumorigenesis). It is also an
CC immunogen for antibody production, as a reagent for detecting hek or hek-
CC L in in vitro assays, to determine binding of hek proteins, to purify hek
CC proteins, and to carry diagnostic or cytotoxic agents to particular
CC leukaemia cells that express the hek antigen. Hek-L also binds the elk
CC tyrosine kinase receptors. See also AAR71481. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 201 AA;

Query Match 100.0%; Score 989; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.4e-102;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRLLPLLRTVLWAAFLGSPLRGGSSLRHVVYWNSSNPRLLRGDAVVELGLNDYLDIVCPH 60

Qy         61 YEGPGPPEGPETPALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YEGPGPPEGPETPALYMVDWPGYESCQAEGPRAYKRWVCSLPFGHVQFSEKIQRFTPFSL 120

Qy        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GFEFLPGETYYYISVPTPESSGQCLRLQVSVCCKERKSESAHPVGSPGESGTSGWRGGD 179
```

Search completed: April 19, 2005, 23:13:33
Job time : 48.9895 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 09:30:04 ; Search time 344.062 Seconds
(without alignments)
8258.620 Million cell updates/sec

Title: US-09-904-954-3_COPY_94_573
Perfect score: 480
Sequence: 1 GGCTCCAGCCTCCGCCACGT.....GAGGGGGGGACACTCCCAGC 480

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 89%
Maximum Match 100%
Listing first 1500 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	480	100.0	636	2	AAQ85888	Aaq85888 Human hek
2	480	100.0	1181	8	ABZ34863	Abz34863 Coding se
3	478.4	99.7	606	6	ABV78135	Abv78135 Human eph
4	478.4	99.7	606	6	ABZ35711	Abz35711 Human eph
5	478.4	99.7	606	6	ABX09954	Abx09954 Human eph
6	478.4	99.7	606	6	ABL91676	Abl91676 Human pol
7	432.8	90.2	770	12	ADN02769	Adn02769 Human rec

ALIGNMENTS

RESULT 1
AAQ85888
ID AAQ85888 standard; cDNA to mRNA; 636 BP.
XX
AC AAQ85888;
XX
DT 25-MAR-2003 (revised)
DT 03-OCT-1995 (first entry)
XX
DE Human hek-L protein cDNA clone C6.
XX
KW Ligand; cell surface; tyrosine kinase receptor; tumorigenesis; immunogen;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 28. .633
FT /*tag= a
FT sig_peptide 28. .93
FT /*tag= b
FT mat_peptide 94. .630
FT /*tag= c
XX
PN WO9506065-A1.
XX
PD 02-MAR-1995.
XX
PF 17-AUG-1994; 94WO-US009282.
XX
PR 20-AUG-1993; 93US-00109745.
PR 30-AUG-1993; 93US-00114426.
PR 03-DEC-1993; 93US-00161132.
PR 09-MAY-1994; 94US-00240124.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Beckmann MP, Cerretti DP;
XX
DR WPI; 1995-106811/14.
DR P-PSDB; AAR71482.
XX
PT New isolated DNA encoding hek-L protein or its fusion products - useful
PT as assay reagent or for carrying therapeutic and diagnostic compounds to
PT leukaemia cells.
XX
PS Claim 3; Page 37; 45pp; English.
XX
CC The sequence is that of a clone encoding hek-L protein, a protein that
CC can bind hek (a cell surface receptor tyrosine kinase). Hek-L is the
CC first known ligand for hek and can be used to study cellular processes
CC regulated by hek (which may be involved in tumorigenesis). It is also an
CC immunogen for antibody production, as a reagent for detecting hek or hek-
CC L in in vitro assays, to determine binding of hek proteins, to purify hek
CC proteins, and to carry diagnostic or cytotoxic agents to particular
CC leukaemia cells that express the hek antigen. Hek-L also binds the elk
CC tyrosine kinase receptors. See also AAQ85887. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 636 BP; 102 A; 202 C; 186 G; 146 T; 0 U; 0 Other;

Query Match 100.0%; Score 480; DB 2; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.2e-129;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 153

Qy         61 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 213

Qy        121 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 273

Qy        181 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 333
```

```
Qy        241 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300
Db        334 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300

Qy        301 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360
Db        394 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360

Qy        361 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420
Db        454 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420

Qy        421 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 513
Db        514 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 480

Qy        121 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 180
Db        214 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 180

Qy        181 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 240
Db        274 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 240

Qy        241 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300
Db        334 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300

Qy        301 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360
Db        394 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360

Qy        361 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420
Db        454 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420

Qy        421 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 513
Db        514 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 480
```

RESULT 2
ABZ34863
ID ABZ34863 standard; cDNA; 1181 BP.
XX
AC ABZ34863;
XX
DT 04-FEB-2003 (first entry)
XX
DE Coding sequence SEQ ID 221, differentially expressed in osteogenesis.
XX
KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
KW osteoporosis; bone disease; upregulator; human; ephin-ephrin; ss.
XX
OS Homo sapiens.
XX
PN WO200281745-A2.
XX
PD 17-OCT-2002.
XX
PF 05-APR-2002; 2002WO-IB002211.
XX
PR 05-APR-2001; 2001US-0281400P.
XX
PA (AVET) AVENTIS PHARMA SA.
XX
PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
PI Connolly T, Jackson A, Bushnell SE, Rawadi G;
XX
DR WPI; 2003-058567/05.
XX
PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
PT useful for bone disease therapy in subject.
XX
PS Claim 1; Page 214; 237pp; English.
XX
CC The present invention relates to novel nucleotide sequences, which are
CC differentially expressed in models of osteogenesis upon being put in
CC contact with a stimulator of osteogenesis. The present sequence is one
CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
CC disease in a patient, promoting osteogenesis and/or preventing
CC osteoporosis/bone disease. The present sequence encodes a Ephin-ephrin
CC family protein
XX
SQ Sequence 1181 BP; 234 A; 353 C; 331 G; 263 T; 0 U; 0 Other;

Query Match   100.0%; Score 480; DB 8; Length 1181;
Best Local Similarity   100.0%; Pred. No. 2.7e-129;
Matches 480; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 60
Db         94 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 153

Qy         61 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 120
Db        154 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 213
```

RESULT 3
ABV78135
ID ABV78135 standard; DNA; 606 BP.
XX
AC ABV78135;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human ephrin-A3 DNA SEQ ID NO 19.
XX
KW RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic;
KW virucide; protozoacide; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200255693-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-EP000152.
XX
PR 09-JAN-2001; 2001DE-01000586.
PR 26-OCT-2001; 2001DE-01055280.
PR 29-NOV-2001; 2001DE-01058411.
PR 07-DEC-2001; 2001DE-01060151.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-590671/63.
XX
PT Inhibiting expression of target gene, useful e.g. for inhibiting
PT oncogenes, by administering double-stranded RNA complementary to the
PT target and having an overhang.
XX
PS Claim 10; Page 123; 203pp; German.
XX
CC The invention relates to inhibiting expression of a target gene (I) in a
CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
CC structure of at most 49 consecutive bases. At least part of one strand
CC (as1) of dsRNA1 is complementary to (I) and at least one end of dsRNA1
CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC in humans, also genes in Plasmodium or in viruses or viroids that are
CC pathogenic for humans, animals or plants. Introducing an overhang into
CC dsRNA greatly increases effectiveness for inhibiting gene expression,

```
CC   both in vivo and in vitro and also increases stability and thus the
CC   effective concentration inside the cell. The present sequence is that of
CC   a gene related to the invention
XX
SQ   Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

  Query Match             99.7%;  Score 478.4;  DB 6;  Length 606;
  Best Local Similarity   99.8%;  Pred. No. 6.4e-129;
  Matches  479;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 126

Qy         61 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 186

Qy        121 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 246

Qy        181 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 306

Qy        241 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        307 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTCGGCTTT 366

Qy        301 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 426

Qy        361 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 486

Qy        421 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 546


RESULT 4
ABZ35711
ID   ABZ35711 standard; DNA; 606 BP.
XX
AC   ABZ35711;
XX
DT   07-FEB-2003  (first entry)
XX
DE   Human ephrin A3 encoding polynucleotide SEQ ID NO 19.
XX
KW   Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW   protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
KW   virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW   Hepatitis C virus; human papilloma virus; gene; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100588-A1.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2001; 2001DE-01000588.
XX
PR   09-JAN-2001; 2001DE-01000588.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R,  Limmer S,  Rost S,  Hadwiger P;
XX
DR   WPI; 2002-683450/74.
XX
PT   Inhibiting expression of target genes, useful e.g. for treating tumors,
PT   by introducing into cells two double-stranded RNAs that are complementary
PT   to the target.
XX
PS   Claim 13; Page 19-20; 100pp; German.
XX
CC   The invention relates to inhibiting expression of a target gene in a cell
CC   by introducing at least two oligoribonucleotides (dsRNAI and II), both
CC   with a double-stranded (ds) structure of at most 49 sequential nucleotide
CC   pairs. At least part of one strand (S1, S2) of the ds structures in each
CC   of dsRNAI and II are complementary to regions in the target gene. The
CC   method uses antisense inhibition of gene expression using double stranded
CC   RNA inhibition (RNAi). The method is particularly used to treat tumours
CC   or infections, especially by Plasmodium or viruses/viroids (pathogenic on
CC   humans, animals or plants). The method provides more effective inhibition
CC   of expression than known methods using a single dsRNA, even at very low
CC   concentrations. When dsRNA has at least one unpaired nucleotide at the
CC   end, stability (and thus effective concentration in the cell) is improved
CC   and efficiency can be increased further by pretreating the cells with
CC   interferon. The present sequence is that of a target DNA of the invention
XX
SQ   Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

  Query Match             99.7%;  Score 478.4;  DB 6;  Length 606;
  Best Local Similarity   99.8%;  Pred. No. 6.4e-129;
  Matches  479;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 126

Qy         61 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 186

Qy        121 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 246

Qy        181 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 306

Qy        241 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        307 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTCGGCTTT 366

Qy        301 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 426

Qy        361 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 486

Qy        421 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 546


RESULT 5
ABX09954
ID   ABX09954 standard; DNA; 606 BP.
XX
AC   ABX09954;
XX
DT   23-JAN-2003  (first entry)
XX
DE   Human ephrin A3 DNA fragment SEQ ID 19.
XX
KW   Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW   prion; inhibition; human; ds.
```

XX
OS Homo sapiens.
XX
PN DE10100587-C1.
XX
PD 21-NOV-2002.
XX
PF 09-JAN-2001; 2001DE-01000587.
XX
PR 09-JAN-2001; 2001DE-01000587.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-742209/81.
XX
PT Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT introduction of complementary double-stranded oligoribonucleotide, after
PT treating the cell with interferon.
XX
PS Disclosure; Page 24-25; 98pp; German.
XX
CC This invention describes a novel method for inhibiting expression of a
CC target gene by introducing into the cell that contains the target gene at
CC least one oligoribonucleotide (dsRNAI) that has a double-stranded (ds)
CC structure of not more than 49 consecutive nucleotides (nt), where at
CC least a segment of one strand of the ds structure is complementary with
CC the target gene and the cells are treated with interferon before
CC introduction of dsRNAI. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX09936-ABX10075 represent
CC gene fragments used to illustrate the method of the invention
XX
SQ Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

Query Match 99.7%; Score 478.4; DB 6; Length 606;
Best Local Similarity 99.8%; Pred. No. 6.4e-129;
Matches 479; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 126

Qy         61 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 186

Qy        121 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 246

Qy        181 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 306

Qy        241 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        307 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTCGGCTTT 366

Qy        301 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 426

Qy        361 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 486
```

```
Qy        421 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 546
```

RESULT 6
ABL91676
ID ABL91676 standard; DNA; 606 BP.
XX
AC ABL91676;
XX
DT 28-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 19.
XX
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS Homo sapiens.
XX
PN DE10100586-C1.
XX
PD 11-APR-2002.
XX
PF 09-JAN-2001; 2001DE-01000586.
XX
PR 09-JAN-2001; 2001DE-01000586.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-270454/32.
XX
PT Inhibiting gene expression in cells, useful for e.g. treating tumors, by
PT introducing double-stranded complementary oligoRNA having unpaired
PT terminal bases.
XX
PS Claim 13; Page 20-21; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a single-
CC stranded segment of 1-4 nt. The method provides oligoribonucleotides for
CC antisense inhibition of gene expression useful e.g. for treating tumours
CC but the oligoribonucleotides may also be directed against genes present
CC in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans,
CC animals or plants) or against cytokine, Id, developmental or prion genes.
CC The method provides more effective inhibition of gene expression than use
CC of known oligonucleotides, probably because the unpaired overhang
CC increases stability and thus intracellular concentration
XX
SQ Sequence 606 BP; 96 A; 191 C; 175 G; 144 T; 0 U; 0 Other;

Query Match 99.7%; Score 478.4; DB 6; Length 606;
Best Local Similarity 99.8%; Pred. No. 6.4e-129;
Matches 479; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GGCTCCAGCCTCCGCCACGTAGTCTACTGGAACTCCAGTAACCCCAGGTTGCTTCGAGGA 126

Qy         61 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTAGACATTGTCTGCCCCCACTACGAA 186

Qy        121 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCTTTGTACATGGTGGACTGGCCAGGC 246
```

```
Qy        181 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTACAAGCGCTGGGTGTGCTCCCTGCCC 306

Qy        241 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCTTTCTCCCTCGGCTTT 300
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        307 TTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGCTTCACACCCTTCTCCCTCGGCTTT 366

Qy        301 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGTTCTTACCTGGAGAGACTTACTACTACATCTCGGTGCCCACTCCAGAGAGTTCTGGC 426

Qy        361 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAGGAGAGGAAGTCTGAGTCAGCCCAT 486

Qy        421 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGGTGGCGAGGGGGGGACACTCCCAGC 546
```

```
RESULT 7
ADN02769
ID   ADN02769 standard; cDNA; 770 BP.
XX
AC   ADN02769;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human receptor and membrane-associated protein cDNA #22.
XX
KW   ds; gene; cytostatic; antiarteriosclerotic; anti-HIV; cerebroprotective;
KW   antiparkinsonian; nootropic; neuroprotective; immunosuppressive;
KW   antiinflammatory; antiallergic; anabolic; hypertensive; anorectic;
KW   antidiabetic; gene therapy; receptor and membrane-associated protein;
KW   REMAP; diagnosing; cancer; atherosclerosis; AIDS; allergy;
KW   Parkinson's disease; Alzheimer's disease; stroke; Addison's disease;
KW   obesity; diabetes; microarray; gene expression; receptor.
XX
OS   Homo sapiens.
XX
PN   WO2004029218-A2.
XX
PD   08-APR-2004.
XX
PF   26-SEP-2003; 2003WO-US030894.
XX
PR   27-SEP-2002; 2002US-0414302P.
PR   11-OCT-2002; 2002US-0417797P.
PR   16-OCT-2002; 2002US-0419217P.
PR   29-OCT-2002; 2002US-0422375P.
PR   07-NOV-2002; 2002US-0424799P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Marquis JP,  Tran UK,  Lee SY,  Richardson TW,  Chawla NK;
PI   Hafalia AJA,  Becha SD,  Ramkumar J,  Khare R,  Tang YT,  Yue H;
PI   Baughn MR,  Elliott VS,  Swarnakar A,  Lu DAM,  Policky JL;
PI   Thangavelu K,  Gietzen KJ,  Blake JJ,  Ison CH;
XX
DR   WPI; 2004-316100/29.
DR   P-PSDB; ADN02719.
XX
PT   New human receptors and membrane-associated proteins and polynucleotides
PT   for diagnosing, preventing or treating diseases associated with aberrant
PT   protein expression, e.g. cancer, atherosclerosis, AIDS, stroke or
PT   diabetes.
XX
PS   Claim 5; SEQ ID NO 72; 268pp; English.
XX
CC   The invention relates to novel human receptor and membrane-associated
CC   protein (REMAP) and the genes encoding them, a naturally-occurring amino
CC   acid sequence that is at least 90% to at least 98% identical any of the
CC   amino acid sequences cited above; or a biologically active or immunogenic
CC   fragment of the polypeptide. The specification also discloses a naturally
CC   occurring polynucleotide sequence that is at least 90% to at least 99%
CC   identical to any of the nucleotide sequences cited above; their
CC   complements or an RNA equivalent. The composition and methods are useful
CC   for diagnosing, preventing or treating diseases or conditions associated
CC   with aberrant expression of REMAP, such as cell proliferative (e.g.
CC   cancer or atherosclerosis), autoimmune/inflammatory (e.g. AIDS or
CC   allergies), neurological (e.g. Parkinson's disease, Alzheimer's disease
CC   or stroke), metabolic (e.g. Addison's disease or obesity), developmental
CC   or endocrine (e.g. diabetes) disorders. These may also be used for
CC   assessing the effects of exogenous compounds on the expression of nucleic
CC   acid and amino acid sequences of REMAP. The REMAP or its fragments are
CC   also useful in screening compounds for effectiveness as agonist or
CC   antagonist of the polypeptides, or in altering the expression of the
CC   target polynucleotide and compounds that specifically bind to or modulate
CC   the activity of the polypeptide. The microarray is useful in monitoring
CC   or measuring protein-protein interactions, drug-target interactions, and
CC   gene expression profiles. This sequence corresponds to the cDNA encoding
CC   one of the proteins of the invention.
XX
SQ   Sequence 770 BP; 134 A; 242 C; 213 G; 181 T; 0 U; 0 Other;

  Query Match             90.2%;  Score 432.8;  DB 12;  Length 770;
  Best Local Similarity   98.4%;  Pred. No. 1.3e-115;
  Matches  437;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;
```

```
Qy         37 AGTAACCCCAGGTTGCTTCGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTA 96
              |||    | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 AGTCTGGCAGGGTTGCTTCGAGGAGACGCCGTGGTGGAGCTGGGCCTCAACGATTACCTA 311

Qy         97 GACATTGTCTGCCCCCACTACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCT 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GACATTGTCTGCCCCCACTACGAAGGCCCAGGGCCCCCTGAGGGCCCCGAGACGTTTGCT 371

Qy        157 TTGTACATGGTGGACTGGCCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTAC 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTGTACATGGTGGACTGGCCAGGCTATGAGTCCTGCCAGGCAGAGGGCCCCCGGGCCTAC 431

Qy        217 AAGCGCTGGGTGTGCTCCCTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 AAGCGCTGGGTGTGCTCCCTGCCCTTTGGCCATGTTCAATTCTCAGAGAAGATTCAGCGC 491

Qy        277 TTCACACCTTTCTCCCTCGGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCG 336
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TTCACACCCTTCTCCCTCGGCTTTGAGTTCTTACCTGGAGAGACTTACTACTACATCTCG 551

Qy        337 GTGCCCACTCCAGAGAGTTCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAG 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GTGCCCACTCCAGAGAGTTCTGGCCAGTGCTTGAGGCTCCAGGTGTCTGTCTGCTGCAAG 611

Qy        397 GAGAGGAAGTCTGAGTCAGCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGG 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GAGAGGAAGTCTGAGTCAGCCCATCCTGTTGGGAGCCCTGGAGAGAGTGGCACATCAGGG 671

Qy        457 TGGCGAGGGGGGGACACTCCCAGC 480
              ||||||||||||||||||||||||
Db        672 TGGCGAGGGGGGGACACTCCCAGC 695
```

```
Search completed: April 19, 2005, 16:26:11
Job time : 345.205 secs
```